Estimation of Theorectically Plausible Demand Functions from U.S. Consumer Expenditure Survey Data

Lester D. Taylor

University of Arizona

Abstract

Estimation of Theoretically Plausible Demand Functions
From U. S. Consumer Expenditure Survey Data

Lester D. Taylor
University of Arizona

The purpose of this paper is the application of four popular theoretically plausible consumer demand systems to a common cross-sectional data set that combines expenditure data from the quarterly BLS consumer expenditure surveys with price data that are collected quarterly in cost-of-living surveys conducted by ACCRA. Six broad categories of expenditure that exhaust total expenditure are analyzed: food consumed at home, housing, utilities, transportation, health care, and miscellaneous. The four demand systems investigated are the Almost-Ideal-Demand-System, the Linear Expenditure System, and the Indirect and Direct Addilog models. Despite absolute differences in magnitudes that in some instances are rather large, there is substantial agreement in the rank-orderings of elasticities. In general, the largest elasticities (for both own-price and total expenditure) are for transportation, miscellaneous, and housing expenditures, while the smallest elasticities (again for both own-price and total expenditure) are for food and utility expenditures. Engel's Law for food is confirmed in all instances.

Estimation of Theoretically Plausible Demand Functions
From U. S. Consumer Expenditure Survey Data

Lester D. Taylor
University of Arizona

I. Introduction

The estimation of demand systems in which the demand equations satisfy all of the restrictions of neo-classical theory of demand has long been considered a triumph of applied econometrics, for the resulting equations both honor the budget constraint and are consistent with an underlying (usually ordinal) utility function. Many such "theoretically plausible" systems engage the literature, ranging from simple linear expenditure equations to systems embodying exotic Gorman Polar Forms. The purpose of this paper is to report results from applying four of the more popular of these systems to data sets that combine household expenditure data from the quarterly Consumer Expenditure Surveys quarterly by the Bureau of Labor Statistics with price data obtained in quarterly surveys conducted by ACCRA. The four systems investigated are: the Almost Ideal Demand System of Deaton and Muellbauer (1980), the Linear Expenditure System of Stone (1954), and the Indirect and Direct Addilog systems of Houthakker (1960). Most of the focus in the paper is on two things: the ease (or lack thereof) with which the theoretically plausible systems can be implemented and estimated, and comparison of the price and total expenditure elasticities that are obtained.

II. Theoretical Considerations

Although the theory of consumer choice is one of the best-developed branches of economic theory, and its (often elegant) presentation is available in many places, it will be nevertheless be useful for present purposes to begin with a brief overview of the framework that is involved. An economic agent, identified as an individual consumer, is assumed to allocate an income of y over n market goods q_i, which can be purchased at unit prices of p_i, in such a way that a "utility" function defined over the n goods, $\varphi(q_1, ...,q_n)$, is at a maximum. More formally, purchase decisions are assumed to follow as the solution to the following constrained maximization problem:

Find the values of q_i, i = 1, ..., n, that maximizes the function,

$$(1) \qquad \varphi(q) \quad = \quad \varphi(q_1, ...,q_n),$$

subject to the condition that

$$(2) \qquad \sum p_i q_i \quad = \quad y.$$

To solve this problem, one first formulates the expression,

$$(3) \qquad \Phi(q, \lambda) \quad = \quad \varphi(q) - \lambda(y - \sum p_i q_i),$$

where λ is a Lagrangean multiplier representing the marginal utility of income, differentiates this expression with respect to the q_i and λ:

(4) $\partial\Phi/\partial q_i$ = $\partial\varphi/\partial q_i$ - λp_i , $i = 1, ... , n,$

(5) $\partial\Phi/\partial\lambda$ = y - $\sum p_i q_i$,

equates the n + 1 derivatives to zero, and then solves the resulting first-order conditions for the n demand functions q_i as functions of the n prices p_i and income y:[1]

(6) q_i = $q_i(p_1, ...,p_n , y),$ $i = 1, ... , n.$

The explicit expressions for the demand functions in (6) obviously depends upon the analytic form of the utility function. For some utility functions, such as the Stone-Geary utiltiy function (which yields the equations associated with the Linear Expenditure System), the demand functions are easily derived and often fairly easily estimated, while for other utility functions (such as the direct addilog utility function of Houthakker), the demand functions are both complicated and highly non-linear. In the latter case, estimation of the parameters necessary for calculating price and income elasticities can sometimes proceed via iteration on λ in expression (4). An instance of this for the direct addilog model will be given below.[2]

III. Background and Merging of Data Sets

Household budget surveys have had a variety of uses in a long and venerable history, ranging from concern with the "state of the poor" in late 18th Century and mid-19th Century England and Continental Europe to a need for weights to be used in construction of consumer price indices.[3] For economists, the principal use of data from household budget surveys has usually been in the analysis of relationships between consumption expenditures and income [i.e., in the analysis of what, since

---

[1] Solution of the first-order conditions requires that the utility function φ(q) satisfy a variety of regularity conditions. For present purposes, we will take these to include φ to be continuous in the n q_i , with continuous first and second partial derivatives that are positive and negative, respectively.

[2] Estimation of demand systems by iterating on the marginal utility of income was first undertaken, as far as I am aware in the context of a quadratic utility by Houthakker and Taylor (1970). Cf., also Taylor and Weiserbs (1972).

[3] The standard reference for the history of early empirical studies of consumer behavior using data from household budget surveys is Stigler (1954); see also Houthakker (1957). Important 20th century studies with a family budget focus include Allen and Bowley (1935), Shultz (1938), Prais and Houthakker (1955), Deaton and Muelbauer (1980), and Pollak and Wales (1992).

Engel (1857), have been known as *Engel Curves*]. Since most budget surveys collect only expenditure data, rather than both quantities and prices, it is generally not possible, absent heroic theoretical assumptions on the structure of consumer preferences, to estimate full-blown demand functions, and hence to obtain estimates of both income and price elasticities.[4]

In the absence of information on prices, estimation of demand functions using expenditure data obviously requires price data from some other source. For the BLS-CES surveys, the natural place to turn for such data is in the price surveys that the Bureau of Labor Statistics pursues monthly as input into construction of the Consumer Price Indices. Prices for several hundred categories of expenditure for some 140 urban areas are collected in these surveys, so that cross-sectional price variation is in principle available. However, the problem is that indices reflecting areal variation in price levels at a point in time are not currently constructed by BLS, but rather only indices that measure price variation over time. Thus, the fact that the BLS all-items index for October, 2003, is 190.3 for Philadelphia and 196.3 for San Francisco *cannot* be interpreted as saying that the all-items CPI was 1.03 percent higher in San Francisco than in Philadelphia, but only that the all-items index in Philadelphia was 190.3 percent higher in October, 2003, than it was during the base years of 1982-1984, and similarly for San Francisco. Thus, the areal price indices that are currently constructed by BLS unfortunately cannot serve the need at hand.

A second source of price information is in surveys that are conducted quarterly by ACCRA in 320 or so U. S. cities.[5] Prices are collected by ACCRA for about 60 items of consumption expenditure, from which city-specific indices can be constructed that can be used to measure price differences both though time for a specific city and across cities at a point in time. In principle, this is precisely the form of price information that is required. From the 60 or so items for which price data are collected, ACCRA constructs indices for six broad categories of expenditure, namely, groceries, housing, utilities, transportation, health care, and miscellaneous. The items underlying the six ACCRA categories are given in Part A of the appendix.

In the analyses to follow, the six ACCRA categories are allied with comparable categories in the BLS CES surveys. In particular, the ACCRA category "groceries" is identified with the CES category "food consumed at home", while the other four specific ACCRA categories are identified with CES counterparts of the same name. Finally, the ACCRA miscellaneous category is identified with CES total expenditure minus the sum of expenditures for the first five categories. Since, to protect confidentiality, place of residence in the CES samples is specified only in terms of state and size of urban area, the ACCRA city price indices have had to be aggregated to a state level. Weights

---

[4] The reference here is to price elasticities estimated from conventional household budget surveys. Deaton (1997) provides an exception. In contrast, estimation of price elasticities for goods, such as telephone or utility services, in which the data used in estimation are collected from the records of vendors or from the actual bills of consumers are fairly commonplace. Cf., Taylor and Kridel (1993) and Rappoport and Taylor (1997).

[5] See www.ACCRA.com.

used in the aggregation are city population from the U. S. Census of 2000. The resulting state-level price indices are then attached to households in the CES samples according to states of residence.

While attaching prices from ACCRA surveys to the CES samples in the manner described yields a cross-sectional consumption data set in which both price and income elasticities can be estimated, it is important to keep in mind that *any* attempt to extract price elasticities from household budget data, not just the present effort, is laden with difficulties. The easiest case, of course, is where a good is both narrowly defined and homogeneous, and the price variation is due solely to price differences between regions. In this circumstance, the problem is simply one of obtaining an appropriate set of prices. With non-homogeneous goods, on the other hand, the situation is much more complicated. For not only does price become ambiguous, but so too does the concept of quantity. Quality differences, which are almost always present in some degree in consumer expenditure data, are especially troublesome in this regard, as is also non-homogeneity arising from broad categories of goods. Not surprisingly, both problems have attracted a great deal of attention in the literature.[6] Finally, a third form of price variation that warrants consideration is that caused by regional differences in the cost-of-living. A haircut, for example, may be more expensive in New York City than in Wichita, in part because of scarcity, but in part also because of differences in the cost-of-living.

As noted, the ideal circumstance (at least in principle) is where goods are narrowly defined and homogeneous (i.e., no grouping or quality gradations), and the price variation is due entirely to different prices for the same good (i.e., no cost-of-living effects). The task in this situation is simply to match expenditures for each household with the prices that the households paid. Since expenditure is quantity times price, it obviously does not matter whether consumption is measured in terms of quantity or expenditure. Price and expenditure elasticities can be translated into one another through the addition or subtraction of 1. Unfortunately, however, the ideal circumstance just described is obviously not the one at hand. Consumption categories in the CES surveys are not narrowly defined, quality gradations are almost certainly present, and the same is true of regional differences in cost-of-living. While efforts are made in the presentation to follow to mitigate the problems that these lapses entail, notions that the price elasticities obtained are the clean, pristine ones of theory are best put in abeyance.

IV. Empirical Results

We now turn to the empirical results of the exercise, which involves the estimation of four separate theoretically plausible systems of demand equations using a combined CES-ACCRA data set for the four quarters of 1996. In estimation of each system, the procedure has been to employ the same set of socio-demographical-regional variables as additional predictors. These variables, which include such quantities as household size, age, and education, are included (always in the same form) in each demand equation in order to control for variations in socio-demographical and regional

---

[6] The discussion of these problems in Prais and Houthakker (1955) is as fresh today as when it was first written 50 years ago.

characteristics across households. A full listing of the variables in question is given in Part B of the appendix. However, since, for present purposes, the focus is on the estimation of price and income elasticities, these other variables will be ignored.[7]

1. The Almost Ideal Demand System

Since its introduction by Deaton and Muellbauer in 1980, the Almost Ideal Demand System (AIDS) has been a workhorse of applied demand analysis, not least because of the ease with which its demand equations can be estimated. The estimating equations of the Deaton-Muellbauer system have the form:

$$(7) \quad w_i \quad = \quad \alpha_i + \sum \gamma_{ij} \ln p_j + \beta_i \ln(y/P), \quad i,j = 1, ..., n,$$

where w_i denotes the budget share of the ith good, p_j denotes the price of the jth good, y denotes the budget constraint, and P is a price index defined by:[8]

$$(8) \quad \ln P \quad = \quad \alpha_0 + \sum \alpha_j \ln p_j + (1/2) \sum \sum \gamma_{ij} \ln p_j p_i .$$

The matrix of own-and cross-price and total-expenditure elasticities obtained from estimating this system of equations for the six BLS-ACCRA expenditure categories for the four quarters of 1996 are

_____

[7] A perennial question in the analysis of budget surveys is the extent to which dynamics are reflected in expenditure data. If dynamics are absent, then the price and income elasticities that are being estimated here can be interpreted as measuring long-run (or steady-state) values, whereas if dynamics are present the estimates are neither fish nor fowl, in the sense of being neither short-run nor long-run. Following a debate in the 1950's concerning the efficacy of incorporating income elasticities that are extraneously estimated from budget surveys into time-series regressions for estimating price elasticities, the view has pretty much been that the situation with budget data is the former, that is, that short-term dynamics are largely absent, so that the estimates obtained (assuming that models are otherwise properly specified) represent steady-state values. The basis for this argument is that, whereas time-series estimates of price and income elasticities will reflect short-run adjustment to changes in income and prices, cross-section estimates will reflect long-run, steady-state adjustment. The latter is seen as being the case if households, even though they may be in temporal disequilibrium, are affected equally by cyclical and other time-varying factors. For present purposes, the view taken is that the elasticities estimated represent steady-state values.

[8] Since saving is not included as an "expenditure" category, the budget constraint in Table 1 (as well as the tables to follow) is total expenditure (defined as the sum of expenditures for food consumed at home, housing, utilities, transportation, and miscellaneous expenditures), rather than CES after-tax income. To simplify estimation, I have used the ACCRA all-items index in place of P as defined in expression (8). Others' experience in estimating the Deaton-Muellbauer system suggests that any bias that this might cause should not be large.

Table 1

Price and Total Expenditure Elasticities
Almost Ideal Demand System
CES-ACCRA Surveys 1996

(calculated at sample mean values)

	food	shelter	utilities	trans.	healthcare	misc.	total expenditure
food	-0.2981	0.6644	0.0599	-0.0013	0.1400	-0.5044	0.4469
shelter	-0.1105	-0.8285	0.1909	0.1902	0.2782	-0.5777	0.8876
utilities	-0.1071	0.1638	-0.7222	0.0523	-0.0669	0.1783	0.4612
trans.	-0.6134	-0.2520	-0.2471	-1.3739	-0.7627	1.5824	1.7250
healthcare	-0.7813	0.0023	0.4260	-0.0129	-0.9375	0.8318	0.6338
misc.	0.4395	-0.2179	-0.2267	-0.0154	0.0470	-1.1448	1.2150

tabulated in Table1.[9] Discussion of the numbers in this and the tables that follow is postponed until Section V.

2. The Linear Expenditure System

The Linear Expenditure System has its theoretical basis in the Stone-Geary-Samuelson utility function, which has the form:

(9) $\quad \varphi(q) = \sum \beta_i \ln(q_j - \alpha_i), \qquad i = 1, ..., n,$

---

[9] The elasticities for the Almost Ideal Demand System models are calculated (at sample mean values) according to the following formulae:

$$\eta_{tot.exp.} = 1 + \beta_i/w_i$$

$$\eta_{ownprice} = -1 + (\gamma_{ii}/w_i) - [\beta_i p_i w_i*]/Pw_i]$$

$$\eta_{cross-price} = (\gamma_{ij}/w_i) - \beta_i p_j w_j*]/Pw_i,$$

where w_i* is the weight of the ith expenditure category in the ACCRA all-items index.

where the α_i have a standard interpretation of "minimum required quantities", and the β_i are subject to the constraint that they sum to 1. The demand functions corresponding to expression (6) are accordingly:

$$(10) \qquad q_j = \alpha_j + \frac{\beta_j(y - \sum p_i \alpha_i)}{p_j} \quad , \qquad j = 1, ..., n.$$

Multiplication by p_j then yields the equations that give the Linear Expenditure System its name:

$$(11) \qquad p_j q_j = p_j \alpha_j + \beta_j(y - \sum p_i \alpha_i), \qquad j = 1, ..., n.$$

Expenditures on the commodities in the Linear Expenditure System are thus seen to consist of amounts $p_j\alpha_j$ that are independent of income, plus proportions βj of the "uncommitted" (or "supernumerary") income, $y - \sum p_i\alpha_i$, that remains. However, since the minimum required expenditures are not observed, they have to be estimated. Estimation has proceeded via an iterative scheme, in which estimates of the α_j's are obtained as the coefficients on p_j in (homogeneous) regressions of $p_j q_j$ on p_j and $y - \sum p_j a_j$, where the a_j's are the estimated α_j's from the preceding iteration. Iteration proceeds until stable estimates of the β_j's are obtained.[10]

The estimated price and total expenditure elasticities for the six CES/ACCRA categories of expenditure are given in Table 2.[11]

---

[10] This estimation scheme differs from the one employed by Stone in his 1954 *Economic Journal* article, in that Stone iterated on β_j rather than on α_j. Iteration on α_j not only appears to be simpler, but convergence to stable estimates of the β_j's is quite rapid The estimates in Table 2 are based upon 10 iterations.

[11] The LES elasticities in this table are calculated according the following formulae:

$$\eta_{tot.exp.} = \beta_j/w_j$$

$$\eta_{ownprice} = -\beta_j(p_j\alpha_j + y - \sum p_i\alpha_i)/p_j q_j$$

$$\eta_{cross-price} = -\beta_j(p_i\alpha_i/p_j\alpha_j).$$

8

Table 2

Price and Total Expenditure Elasticities
Linear Expenditure System
CES-ACCRA Surveys 1996

(calculated at sample mean values)

	food	shelter	utilities	trans.	healthcare	misc.	total expenditure
food	-0.1532	-0.0078	-0.0013	0.0043	0.0005	-0.0047	0.1317
shelter	-0.0156	-0.6074	-0.0052	0.0167	0.0021	-0.0181	0.5902
utilities	-0.0053	-0.0101	-0.1966	0.0057	0.0007	-0.0061	0.1731
trans.	-0.0521	-0.1004	-0.0174	-1.8735	-0.0069	-0.0521	2.6826
healthcare	-0.0086	-0.0167	-0.0029	0.0093	-0.3190	-0.0100	0.2887
misc.	-0.0302	-0.0582	-0.0101	0.0324	0.0040	-1.1536	1.3278

3. The Indirect Addilog Model[12]

In terms of convenience and goodness-of-fit, simple double-logarithmic demand functions are pretty much without peer in applied demand analysis. Nevertheless, problems abound with at the theoretical level, for double-logarithmic demand functions are neither integrable nor additive. In recognition of this, Houthakker (1960), in one of the most influential published papers ever in demand theory, introduced two near-logarithmic demand systems that are both additive and consistent with conventional demand theory. The first of these (the *indirect addilog model*) is derivable from an indirect utility function, while the second (the *direct addilog model*) can be derived from a direct utility function. The indirect addilog model is the easier of the two to implement, and will be presented first.

The procedure employed by Houthakker is to transform double-logarithmic demand functions into an additive system via the fact that any non-additive function $\theta_i(y)$ can be made additive by the transformation,

---

[12] In addition to Houthakker (1960), good discussions of the indirect and direct addilog models can be found in Phlips (1983).

$$(12) \qquad g_i(y) = \frac{y\theta_i(y)}{\sum \theta_k(y)} \quad,$$

since $\sum g_i(y) = y$. The application of this transformation to the double-logarithmic function,

$$(13) \qquad p_i q_i \quad = \quad A_i \beta_j y^{\beta_i} p_i^{-\beta_i} \quad , \quad i = 1, ..., n \, ,$$

gives an additive system of functions:[13]

$$(14) \qquad f_j(y,p) = \frac{A_j \beta_j y^{\beta_j} p_j^{-\beta_j}}{\sum A_j \beta_j y^{\beta_j-1} p_j^{-\beta_j}} \quad , \quad j = 1, ..., n.$$

Division of $f_j(y,p)$ by $f_i(y,p)$ then yields:

$$(15) \qquad \frac{p_j q_j}{p_i q_i} = \frac{A_j y^{\beta_j} p_j^{-\beta_j}}{A_i y^{\beta_i} p_i^{-\beta_j}} \quad,$$

which, upon taking logarithms, becomes:

$$(16) \quad \ln q_j - \ln q_i = a_{ij} + (\beta_j + 1)(\ln y - \ln p_j) - (\beta_i + 1)(\ln y - \ln p_i), \quad j = 1, ..., n, \ j \neq i \,.$$

The price and total expenditure elasticities obtained from applying the system in expression (16) to the six BLS-ACCRA categories appear in Table 3.[14]

---

[13] This expression can be obtained more conventionally by applying Roy's Theorem to the indirect utility function:

$$\varphi(y/p) = \sum A_j (y/p_j)^{\beta_j} \,.$$

[14] The total expenditure and price elasticities for the indirect addilog equations are calculated as follows:

$$\eta_{tot.exp.} \quad = \quad 1 + \beta_j - \sum \beta_j w_j$$

$$\eta_{ownprice} \quad = \quad -1 - (1 - w_j)\beta_j$$

$$\eta_{cross-price} \quad = \quad \beta_i w_i.$$

Table 3

Price and Total Expenditure Elasticities
Indirect Addilog Model
CES-ACCRA Surveys 1996

(calculated at sample mean values)

	food	shelter	utilities	trans.	healthcare	misc.	total expenditure
food	-0.1084	-0.1204	-0.1204	-0.1204	-0.1204	0.1204	0.4406
shelter	-0.1656	-0.6647	-0.1656	-0.1656	-0.1656	-0.1656	0.9518
utilities	-0.0729	-0.0729	-0.1530	-0.0729	-0.0729	-0.0729	0.5327
trans.	-0.0126	-0.0126	-0.0126	-0.9338	-0.0126	-0.0126	1.3738
healthcare	-0.0524	-0.0524	-0.0524	-0.0524	-0.2284	-0.0524	0.6287
misc.	-0.0289	-0.0289	-0.0289	-0.0289	-0.0289	-0.9124	1.3361

4. The Direct Addilog Model

The utility function corresponding to Houthakker's direct addilog model is:

$$(17) \qquad \varphi(q) \quad = \quad \sum \alpha_j q_j^{\beta_j} \; .$$

From the first-order conditions (4) and (5), one obtains:

$$(18) \qquad q_j^{\beta_j - 1} = \frac{1}{\alpha_j \beta_j} \lambda p_j \; , \quad j = 1, \, ..., \, n$$

$$(19) \qquad \lambda = \frac{\sum \alpha_j \beta_j q_j^{\beta_j}}{y} \; ,$$

which in turn yield the demand functions:

$$(20) \qquad q_j^{\beta_j - 1} = \frac{y p_j}{\alpha_j \beta_j \sum \alpha_j \beta_j q_j^{\beta_j}} \; , \quad j = 1, \, ..., \, n \; .$$

In view of the severe non-linearity of the equations in (20), the usual procedure (as with the indirect addilog model) preparatory to estimation of the β_j's is to divide $f_j(y, p)$ by $f_i(y, p)$, and taking logarithms, to obtain:

(21) $\qquad (\beta_j - 1)\ln q_j - (\beta_i - 1)\ln q_i = a_{ji} + \ln p_j - \ln p_i, \quad j = 1, ..., n, j \neq i,$

which upon rearrangement and division by $(\beta_j - 1)$ yields:

(22) $\qquad \ln q_j = b_{ji} + b_{1j}\ln q_i + b_{2j}(\ln p_j - \ln p_i), \quad j = 1, ..., n, j \neq i,$

where:

(23) $\qquad b_{1j} = (\beta_i - 1)/(\beta_j - 1)$

(24) $\qquad b_{2j} = 1/(\beta_j - 1).$

From (23) and (24), estimation of the n - 1 equations in expression (22) accordingly requires that the equations be estimated subject to the n - 2 constraints:

(25) $\qquad b_{1j}b_{2k} = b_{2j}b_{1k}, \quad j, k = 1, ..., n, j \neq k \neq i.$

As will be related in the next section, initial efforts to estimate the β_j's on the basis of expression (22) ran into difficulties, and expression (18) has been resorted to instead. In logarithms, expression (18) becomes:

(26) $\qquad \ln q_j = a_j + \dfrac{1}{\beta_j - 1}\ln \lambda p_j, \quad j = 1, ..., n.$

Estimation of the equations in (26) has proceeded via quasi-iteration on λ, with λ approximated by the function,[15]

(27) $\qquad \lambda \cong (\text{total expenditure})^{-\kappa}.$

The price and total expenditure elasticities with a value for κ equal to 0.5 are presented in Table 4.[16]

---

[15] The reason for the modifier *quasi* will be explained in Section V.

[16] The total expenditure and price elasticities in this table are calculated from the following systems of simultaneous equations:

V. Some Technical *Obiter Dicta* Concerning Estimation

Before taking up a comparison and discussion of the results presented in Tables 1 - 4 for the four demand systems, some remarks concerning their estimation are in order.

1. As noted, of the four demand systems analyzed, the AIDS model is by far the easiest to estimate, as the parameters of the system are obtained as the coefficients in the regressions of the budget shares w_j on $\ln p_i$ ($i = 1, ..., n$) and $\ln(y/P)$.

2. For the linear expenditure system, on the other hand, things are not so straightforward, for the minimum required quantities, α_j, are unobserved, and estimation accordingly has to proceed by some sort of iterative scheme. Stone's original procedure was to iterate on the β's, but the procedure here has been to iterate on the α's in expression (11), whereby (beginning with $\alpha = 0$) the α's estimated as the coefficients on p_j in iteration k - 1 are used to construct $y - \sum p_i \alpha_i$ for iteration k.[17]

3. From expression (16), we see that the n - 1 equations comprising the estimating equations for the indirect addilog model have to be estimated subject to the restriction that β_i has the same value in each equation. This has been effected by estimating the equations in expression (16) in a seemingly-unrelated-regressions framework, that is, by "stacking" the observations for the n - 1 equations, and then estimating the equations in expression (16) in a seemingly-unrelated-regressions framework, that is, by "stacking" the observations for the n - 1 equations, and then estimating the resulting combined equation with a common coefficient on $\ln p_i$, but separate coefficients on the $\ln p_j$ (as well as on all other variables).

$$(1 - \beta_j)\partial \ln q_j / \partial \ln y \; + \; \sum w_i \, \beta_i \, \partial \ln q_i / \partial \ln y \; = \; 1 \, , \quad j, i \; = \; 1, ..., n$$

$$(1 - \beta_j)\partial \ln q_j / \partial \ln p_i \; + \; \sum w_i \, \beta_i \, \partial \ln q_i / \partial \ln p_i \; = \; 1 \, , \quad j, i \; = \; 1, ..., n.$$

[17] The elasticities in Table 2 are derived from 50 iterations. Convergence actually is quite rapid, and results for 40 or more iterations are little changed from those for 10 iterations. Still another procedure for estimating the LES model would be to estimate α and β directly from the first-order conditions (4) by iterating on λ, in which case the estimating equations would be given by:

$$q_j \;\; = \;\; \alpha_j \; + (\lambda \, p_j)^{-1} \beta_j \, , \quad j = 1, ..., n,$$

from which estimates of α_j and β_j can be obtained as coefficients in the regression of q_j on a constant and $(\lambda \, p_j)^{-1}$. Since λ for the LES is equal to $1/ (y - \sum p_i \alpha_i)$, a natural value with which to begin the iterations is $\lambda = 1/y$. Interestingly, the minimum-required-quantities for transportation and health care expenditures are both estimated to be negative. An analysis of this situation can be found in Solari (1971).

Table 4

Price and Total Expenditure Elasticities
Direct Addilog Model
CES-ACCRA Surveys 1996

(calculated at sample mean values)

	food	shelter	utilities	trans.	healthcare	misc.	total expenditure
food	-1.0954	0.0261	0.0075	0.5031	0.0309	1.2940	0.4639
shelter	0.0099	-1.1378	0.0075	0.5031	0.0309	1.2940	0.4885
utilities	0.0099	0.0261	-1.1039	0.5031	0.0309	1.2940	0.4664
trans.	0.0099	0.0261	0.0075	-2.8917	0.0309	1.2940	1.4248
healthcare	0.0099	0.0261	0.0075	0.5031	-1.4564	1.2940	0.6242
misc.	0.0099	0.0261	0.0075	0.5031	0.0309	-2.3213	1.5174

4.   Of the four equation systems analyzed, the direct addilog model, despite what would otherwise appear to be a fairly simple utility function, is by far the most problematic to estimate. The demand functions [cf., expression (20)] are intractable empirically, and estimation (as with the indirect addilog model) is usually to proceed in terms of logarithmic deviations from a "left-out" category. In this case, however, complications in the form of non-linear restrictions on the parameters across equations are brought into play. Three approaches have been pursued in estimation. The first, which is the one yielding the elasticities in Table 4, is (as noted) to estimate the parameters directly from the first-order through quasi-iteration on λ. The strength of this approach is that it avoids having to deal explicitly with non-linearities. Its drawback is that, at least in the way that I have implemented it, the procedure is not entirely objective.[18]

---

[18] Implementation of this procedure involves the estimation of the n equations in expression (26) with λ defined as $y^{-\kappa}$. An initial informal search over values of κ suggested values in a range of 0.15 to 0.50. R^2s of the estimated equations for values of κ in this range are relatively stable, as are the implied price and income elasticities. Interestingly, however, values for $\kappa \geq 0.25$ yield several negative β's, which would imply negative marginal utility. Since the latter is theoretically implausible, a value for κ of 0.20 was finally settled on (for which all of the β's are positive).

The second approach tried was to reformulate the equations in expression (22) as:

(28) $\ln q_j = b_{ji} + b_{2j}[(\beta_i - 1) \ln q_i + \ln p_j + \ln p_i], \quad j = 1, ..., n, j \neq i,$

and then iterate on $(\beta_i - 1)$. The results obtained, however, were nonsense.

Finally, the third approach is simply a "brute-force" one of estimating the parameters of the n - 1 equations in expression (22) jointly (i.e., in a seemingly-unrelated-regressions format) subject to the n - 2 non-linear restrictions given in expression (25). This procedure, which entails the estimation of a model with 38,621 observations, 135 variables, and 4 non-linear constraints amongst the coefficients, has been estimated using a non-linear programming algorithm found in SAS.[19]. The elasticities from this estimation are given in Table 5.

VI. Discussion of Results

The four demand systems analyzed in this exercise represent a broad spectrum of utility structures, ranging from the severe separability restrictions of the indirect and direct addilog models to the only very minimal restrictions of the Almost-Ideal-Demand System. The four systems have all been estimated from a common data set, namely, one consisting of 7724 observations from the four quarterly CES surveys for 1996 augmented with price data from the four ACCRA price surveys for that year. To keep the discussion manageable, focus will be restricted to a comparison of price and total expenditure elasticities. Total expenditure elasticities are collected in Table 6, and own-price elasticities in Table 7. As an added measure for comparison, elasticities estimated from simple double-logarithmic equations are included in the tables as well.

For the total expenditure elasticities in Table 6, the first thing to note is that there is strong agreement across the four systems in rank-ordering of relative magnitudes. The rank-ordering is exact for the AIDS and LES models (as well as for the double-log model), and nearly so for the indirect addilog model (with only miscellaneous expenditures and housing switched in order). The expenditure elasticity for transportation is either the largest or second largest in all of the systems, and food, utility expenditures, and health care (in that order) are usually the smallest. Transportation and miscellaneous expenditures are indicated to be luxury goods (i.e., total expenditure elasticities in excess of 1), while housing is seen to be nearly so. Engel's Law for food (i.e., a total expenditure elasticity less than 1) is confirmed in all cases. Finally, it is of interest that the equation system with

---

[19] The algorithm in question is proc nlp in the SAS OR software. Values of b_{1j} and b_{2j} from unconstrained OLS regressions were used as seeds in the constraints. Unfortunately, three of the β's that are obtained turn out to be negative (which, as just noted, imply negative marginal utilities).

15

Table 5

Price and Total Expenditure Elasticities
Direct Addilog Model
CES-ACCRA Surveys 1996
Constrained Non-Linear

(calculated at sample mean values)

	food	shelter	utilities	trans.	healthcare	misc.	total expenditure
food	-0.4573	-0.1969	0.2956	-0.2021	-0.0332	-0.0846	0.6996
shelter	-0.0661	-0.7280	0.2956	-0.2021	-0.2021	-0.0846	0.9498
utilities	-0.0661	-0.0661	-0.4633	-0.2021	-0.2021	-0.0846	1.1546
trans.	-0.0661	-0.0661	0.2956	-1.0729	-0.2021	-0.0846	1.8827
healthcare	-0.0661	-0.0661	0.2956	-0.2021	-1.2384	-0.0846	2.1553
misc.	-0.0661	-0.0661	0.2956	-0.2021	-0.2021	-0.3451	0.4646

Table 6

Total Expenditure Elasticities
AIDS, LES, Indirect, and Direct Addilog Models
CES-ACCRA Surveys 1996

Category	AIDS	LES	Indirect Addilog	Direct Addilog	Double Log	Budget Share
food	0.4469	0.1317	0.5636	0.4639	0.2982	0.1266
shelter	0.8876	0.5902	1.0321	0.4885	0.7826	0.2066
utilities	0.4612	0.1731	0.6401	0.4664	0.3611	0.0897
trans.	1.7250	2.6826	1.5824	1.4248	1.3718	0.1537
healthcare	0.6338	0.2887	0.7440	0.6242	0.4436	0.0703
misc.	1.2150	1.3278	0.9726	1.5174	1.2091	0.3531

total expenditure elasticities closest to those of the double-logarithmic equations is appears to be the LES model.[20]

The own-price elasticities for the four demand systems and double-logarithmic models are tabulated in Table 7. Looking first at the rank-orderings (by relative magnitude) of the elasticities across the models, agreement is again strong, although less so than for the expenditure elasticities. Moreover, the ordering of price elasticities is generally the same as for the total expenditure, transportation and miscellaneous largest, food and utilities smallest, and housing and health care in the middle. Given the strong separability assumptions implicit in LES and the two addilog models (which severely restricts substitution and complementation), strong agreement in relative magnitude between own-price and total expenditure elasticities is of course to be expected.[21] Interestingly, the same agreement is present, though obviously weaker, in the double-log equations (for which the Slutsky-Shultz Relation does not hold).

Turning now to magnitudes, we of course expect the own-price elasticity for food to be small, but ordinarily one does not anticipate an elasticity that is substantially greater than 1 (in absolute value) for transportation, yet such is uniformly the case in all of the models. With regard to the general magnitude of the own-price elasticities, it is to be kept in mind that their default values (their values in the absence of statistical significance) is minus 1 in the AIDS and the two addilog models.[22]

---

[20] The double-log equations are of course not additive in the sense of satisfying the budget constraint. The budget-share weighted sum of the double-log elasticities is 0.87; for the four systems investigated in this exercise, these weighted sums are of course 1 (or nearly so). (The respective budget shares are given in the last column of Table 6.)

[21] This follows from the Slutsky-Shultz Relation [see Wold and Jureen (1953, p.111)] that, since the demand equations for the four demand systems satisfy the budget constraints, the sum of the own-and cross-price elasticities for each expenditure category is equal to the total-expenditure elasticity.

[22] The system that would appear to give the most anomalous results is the direct addilog model, for the fact that all of the own-price elasticities for this model in Table 6 -- even the one for food! -- are in excess of 1 (in absolute value) does not seem plausible. (The expenditure elasticities, on the other hand, are pretty much in line with those in the other models.) However, as noted, the direct addilog model is difficult to estimate, and this, rather than the underlying integrity of the model, may be the problem. The price elasticities from the constrained non-linear estimation in Table 5 seem much more plausible, but (as noted in footnote 19) three of the β's in this estimation are negative.

17

Table 7

Own-Price Elasticities
AIDS, LES, Indirect, and Direct Addilog Models
CES-ACCRA Surveys 1996

Category	AIDS	LES	Indirect Addilog	Direct Addilog	Double Log	Budget Share
food	-0.2981	-0.1532	-0.2277	-1.0954	-0.6663	0.1266
shelter	-0.8285	-0.6075	-0.4846	-1.1378	-0.8198	0.2066
utilities	-0.7222	-0.1966	-0.3440	-1.1039	-0.8806	0.0897
trans.	-1.3739	-1.8735	-1.5260	-2.8917	-1.1990	0.1537
healthcare	-0.9375	-0.3190	-0.4635	-1.4564	-1.2029	0.0703
misc.	-1.1448	-1.1536	-0.5077	-2.3213	-1.1142	0.3531

Hence, although estimates of standard errors have not been calculated, the estimated elasticities that are substantially different than -1 are probably the only ones with any real statistical significance.[23] What is perhaps most surprising about the own-price elasticities, at least to me, are the ones in the double-logarithmic equations, for which the statistical default values are 0 (rather than -1). All are substantial, with those for transportation, health care, and miscellaneous in excess of 1 (in absolute value).[24]

The AIDS model is the only one of the four demand systems that in principle allows for non-income-effect substitution and complementarity, hence it is no surprise that the largest cross-price effects are to be found in Table 1. Since the numbers in the tables are both *elasticities* and based upon *uncompensated* derivatives, for which the Slutsky symmetry conditions do not apply, it is possible for a category (A, say) to be a complement with respect to the price of another good (B, say), but for

[23] See footnotes 8 and 13. Likewise, for these models, the default value for the total-expenditure elasticities is 1.

[24] The t-ratios for the own-price in these equations range from -2.84 and -4.47 for miscellaneous and transportation expenditures to -16.84 and -23.80 for shelter and utilities. The R^2 s for the double-log equations (with 7524 observations) range from 0.1871 for health care to 0.5711 for miscellaneous expenditures. The double-log equations, it should be noted, are estimated with only own-prices in the equations. Inclusion of all of the prices in the equations injects sufficient collinearity into the models for the results not to be meaningful. Obviously, one of the benefits of estimating a demand system is that it provides restraints on multicollinearity.

B to be a substitute with respect to the price of A. Food and housing in Table 1 provide an example of this, for the cross-elasticity of food with respect to the price of housing is -0.11, but the cross-elasticity of housing with respect to the price of food is 0.66. There are many other instances of this switching of signs, so many, in fact, that asymmetry is pretty much the norm.

VII. Conclusions

The motivation for this paper has been the application of four popular theoretically plausible consumer demand systems to a common cross-sectional data set that combines expenditure data from the quarterly BLS consumer expenditure surveys with price data that are collected quarterly in cost-of-living surveys conducted by ACCRA. Six broad categories of expenditure that exhaust total expenditure have been analyzed: food consumed at home, housing, utilities, transportation, health care, and miscellaneous. The four demand systems analyzed are the Almost-Ideal-Demand-System, the Linear Expenditure System, and the Indirect and Direct Addilog models. The focus of the exercise has been on ease (or lack thereof) of estimation and comparison of price and total expenditure elasticities. The AIDS model is the most straightforward to estimate, while the Direct Addilog is the most difficult. Despite absolute differences in magnitudes that in some instances are rather large, there is substantial agreement in the rank-orderings of elasticities. In general, the largest elasticities (for both own-price and total expenditure) are for transportation, miscellaneous, and housing expenditures, while the smallest elasticities (again for both own-price and total expenditure) are for food and utility expenditures. Engel's Law for food is confirmed in all instances. In general, elasticities are smallest with the LES model and largest, perhaps even implausibly so, with the Direct Addilog model.

REFERENCES

Allen, R. G. D. and Bowley, A. L. (1935), *Family Expenditure*, Staples.

Deaton, A. S. (1997), *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Johns Hopkins Press.

Deaton, A. S. and Muellbauer, J. (1980), "An Almost Ideal Demand System," *American Economic Review*, Vol. 70, No. 3, June 1980, pp. 312-326.

Engel, E. (1857), "Die Productions-und Consumptions-verhaltnisse des Konigreichs Sachsen, reprinted in *The Bulletin of the International Statistics Institute*, Vol. 9, No. 1.

Houthakker, H. S. (1957), "An International Comparison of Expenditure Patterns, Commemorating the Centenary of Engel's Law," *Econometrica*, Vol. 25.

Houthakker, H.S. (1960), "Additive Preferences," Econometrica, Vol. 28, pp. 244-257.

Houthakker, H. S. and Taylor, L. D. (1970), *Consumer Demand in the United States (Second Edition)*, Harvard University Press.

Phlips, L. (1983), *Applied Demand Analysis (Second Edition)*, North Holland Publishing Co.

Pollak, R. A. and Wales, T. J. (1992), *Demand System Specification & Estimation*, Oxford University Press.

Prais, S. J. and Houthakker, H. S. (1955), *The Analysis of Family Budgets*, Cambridge University Press.

Rappoport, P. N. and Taylor, L. D. (1997), "Toll Price Elasticities From a Sample of 6500 Residential Telephone Bills," *Information Economics and Policy*, Vol. 9, No. 1, 1997, pp. 51-70.

Schultz, H. (1937), *The Theory and Measurement of Demand*, University of Chicago Press.

Solari, L. (1971), *Theorie des choix et fonctions de consommation semi-agregees*, Droz, Geneva.

Stigler, G. J. (1954), "The Early History of Empirical Studies of Consumer Behavior," *Journal of Political Economy*, Vol. 42, April 1954.

Stone, J.R.N. (1954), "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal*, Vol. 74, No. 255, September 1954, pp. 511-527.

Taylor, L. D. and Kridel, D. J. (1993), "The Demand for Commodity Packages: The Case of Telephone Custom-Calling Features," *The Review of Economics & Statistics*, Vol. 75, No. 2, May 1993.

Wold, H. and Jureen, L. (1953), *Demand Analysis*, John Wiley and Sons.

Appendix

A. Consumption Expenditure Categories Included in ACCRA Price Surveys.

<u>Groceries</u>	<u>Housing</u>	<u>Utilities</u>	<u>Transportation</u>	<u>Health Care</u>	<u>Miscellaneous</u>
t-bone stk.	apt. rent	all electric	bus fare	hosp. room	hamburger sand.
gd. beef	home price	part electric	tire bal.	Dr. appt.	pizza
sausage	mortgage rate	other energy	gasoline	dentist	2-pc. chicken
fry chicken	home P+I	telephone		aspirin	hair cut
tuna					beauty salon
gal. milk					tooth paste
dz. eggs					shampoo
margarine					dry clean
parmesan cheese					men's shirt
potatoes					underwear
bananas					slacks
lettuce					washer repair
bread					newspaper
cigarettes					movie
coffee					bowling
sugar					tennis balls
cereal					monopoly set
sweet peas					liquor
tomatoes					beer
peaches					wine
Kleenex					
Cascade					
Crisco					
orange juice					
frozen corn					
baby food					
Coke					

B. Preparation of Data.

The CES quarterly data sets employed in the analysis have been developed from the Public Use Interview Microdata sets for 1996 that are available on CD-ROM from the U.S. Bureau of Labor Statistics.[25] "Cleansing" of the CES files included elimination of households with reported income of less than $5000 and then of households with zero (or negative) expenditures for the commodity category in question. The CES surveys do not include price data. The price data for the analysis are

---

[25] See http://www.bls.gov/cex/home.htm.

21

taken from the on-going price surveys of the 62 items of consumer expenditure listed in Table A1 above in more than 300 cities in the U.S. that are conducted quarterly by ACCRA[26]. From the 62 items of expenditure, ACCRA constructs six price indices (food, housing, etc.), and then from these an all-items index (which in principle are comparable, on a city basis, to BLS city CPI's). The ACCRA city indices in a state for each quarter are aggregated to the state level using city populations from the US Census of 2000 as weights.[27] The resulting ACCRA prices are then attached to CES households according to state of residence.[28]

C. Definitions of Variables.

lnfood   logarithm of expenditures for food consumed at home

lnhous   logarithm of housing expenditures

lnutil   logarithm of expenditures for household utilities

lntrans   logarithm of transportation expenditures

lnhealth   logarithm of health care expenditures

lnmisc   logarithm of miscellaneous consumption expenditures

lnincome   logarithm of household income

lntotexp   logarithm of total consumption expenditure

lnpfood   logarithm of price index for food consumed at home

lnphous   logarithm of price index for housing

lnputil   logarithm of price index for utility expenditures

lnptrans   logarithm of price index for transportation expenditures

lnphealth   logarithm of price index for health care expenditures

---

[26] See http://www.ACCRA.com.

[27] See http://www.census.gov/Press-Release/www/2003/SF4.html

[28] In instances in which CES does not code state of residence for reasons of non-disclosure, the households in question are dropped.

lnpmisc	logarithm of price index for miscellaneous expenditures
lnpall	logarithm of all-items price index
no_earnr	number of income earners in household
fam_size	size of household
age_ref	age of head of household
dsinglehh	dummy variable for single household
drural	dummy variable for rural area of residence
dnochild	dummy variable for no children in household
dchild1	dummy variable for children in household under age 4
dchild4	dummy variable for oldest child in household between 12 and 17 and at least one child less than 12
ded10	dummy variable for education of head of household: grades 1 through 8
dedless12	dummy variable for education of head of household: some high-school, but no diploma
ded12	dummy variable for education of head of household: high-school diploma
dedsomecoll	dummy variable for education of head of household: some college, but did not graduate
ded15	dummy variable for education of head of household: Bachelor's degree
dedgradschool	dummy variable for education of head of household: post-graduate degree
dnortheast	dummy variable for residence in northeast
dmidwest	dummy variable for residence in midwest

dsouth	dummy variable for residence in south
dwest	dummy variable for residence in west (excluded)
dwhite	dummy variable for white head of household
dblack	dummy variable for black head of household
dmale	dummy variable for male head of household
down	dummy variable for owned home
dfdstmps	dummy variable for household receiving food stamps
D1, D2, D3, D4	seasonal quarterly dummy variables.